REAL-TIME COMPRESSED IMAGING OF SCATTERING VOLUMES

Ohad Menashe, Alexander Bronstein

School of Electrical Engineering
Tel Aviv University

ABSTRACT

We propose a method and a prototype imaging system for real-time reconstruction of volumetric piecewise-smooth scattering media. The volume is illuminated by a sequence of structured binary patterns emitted from a fan beam projector, and the scattered light is collected by a two-dimensional sensor, thus creating an under-complete set of compressed measurements. We show a fixed-complexity and latency reconstruction algorithm capable of estimating the scattering coefficients in real-time. We also show a simple greedy algorithm for learning the optimal illumination patterns. Our results demonstrate faithful reconstruction from highly compressed measurements. Furthermore, a method for compressed registration of the measured volume to a known template is presented, showing excellent alignment with just a single projection. Though our prototype system operates in visible light, the presented methodology is suitable for fast x-ray scattering imaging, in particular in real-time vascular medical imaging.

Index Terms— volumetric reconstruction, structured light, scattering tomography, compressive sensing, sparse coding

1. INTRODUCTION

The overwhelming majority of medical CT images acquired today use attenuation contrast. While allowing to clearly recognize some tissues, this imaging modality does not yield satisfactory sensitivity and specificity for weakly absorbing media, such as various soft tissues [1]. A prominent alternative to transmission x-ray tomography is *scattering tomography*, in which the field of backscattered photons due to Compton scattering is measured instead of the transmitted field [2]. Like the standard transmission CT, scattering tomography requires a large number of measurements in order to reconstruct an accurate image, which extends the acquisition time and prohibits the acquisition of dynamic objects. For appropriate classes of signals, compressed sampling techniques have been demonstrated to defy the Nyquist-Shannon bounds, sometimes by orders of magnitude. Unlike the closed-form reconstruction formula for regularly sampled signals that the Shannon-Nyquist theory offers, reconstruction from compressed samples is more elaborate and requires the solution of an optimization problem, resulting is slow reconstruction speed unsuitable for real time applications. Recent work by [3] followed by [4, 5] and [6, 7] advocate for an alternative approach. These works showed that several established iterative optimization algorithms can be truncated after a small number of iterations and unrolled into a feed-forward architecture resembling a neural network, with the parameters of the layers dictated by the specific functional form of the iteration. While such a truncated algorithm does not fully converge and is incapable to produce satisfactory solutions for *any* measurements vector given as the input, the authors have demonstrated that when the input is restricted to vectors commonly arising from the data distribution, the network parameters can be trained to produce a high-quality approximation to the exact solution.

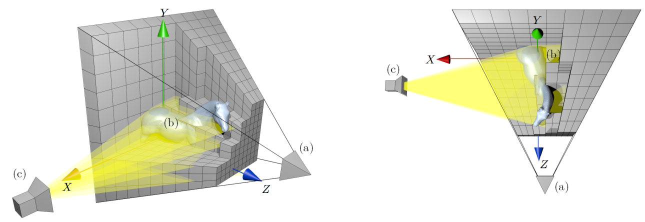

Fig. 1. Scattering tomography imaging system: 3D view(left) and top view (right). The projector(c) illuminates the object(b) with a sequence of one-dimensional patterns, and the scattered light is collected by the sensor(a).

2. SIGNAL FORMATION AND ACQUISITION

We consider a scattering tomography imaging system whose physical setup is depicted schematically in Figure 1. The projector is assumed to admit the standard projective (pinhole) model with the optical axis aligned with the X axis. The sensor, is also assumed to admit a projective model, circular aperture of radius r, and the optical axis aligned with the Z axis. The space in the intersection of the projector and sensor fields of view is assumed to be filled with weakly scattering medium, whose scattering coefficient is denoted by a_{s}.

Standard radiometric calculations lead to the following form of the sensor irradiance

$$E = \pi r^2 \int_0^z I_0 e^{-(\tau_\mathrm{p}+\tau_\mathrm{c})}\, a_\mathrm{s} p(\theta)\, \frac{\cos\phi}{R_\mathrm{c}^2} \frac{\cos\theta}{R_\mathrm{p}^2} \frac{\partial X}{\partial x} J dz, \quad (1)$$

where J denotes the Jacobian of the transformation between the Cartesian and our non-orthogonal coordinates systems, θ is the angle between the normal to the plane $z_p = z$ and the ray $(x_c = x, y_c = y)$, ϕ is the sensor aperture incidence angle, p is the relative density of scattered photons per unit of solid angle, and

$$\tau_\mathrm{p} \;=\; \int_0^x a_\mathrm{s} \frac{\partial X}{\partial x} dx \quad (2)$$

$$\tau_\mathrm{s} \;=\; \int_0^z a_\mathrm{s} \frac{\partial Z}{\partial z} dz \quad (3)$$

denote the optical distances from the projector to the volume element and from the volume element to the sensor aperture, respectively.

However, by assuming the total medium scattering and absorption cross-sections to be sufficiently weak, and adding a weak perspective assumption, equation (1) simplifies to

$$E \;\approx\; \mathrm{const} \cdot \int_0^z I_0 a_\mathrm{s} dz. \quad (4)$$

The latter equation is discretized on a non-Cartesian grid of voxels formed by the x, y, z coordinate

$$E_{ij} \;\approx\; \sum_{k=1}^{N_\mathrm{z}} I_k A_{ijk} w_{ijk}, \quad (5)$$

where E_{ij}, I_k and A_{ijk} denote the discretized versions of $E(x, y)$, $I_0(z)$, and $a_\mathrm{s}(x, y, z)$, respectively, and w_{ijk} are constant weights depending on the system geometry. We note that the discretized expression (5) is separable; for each i, j, we can write $E_{ij} \approx \mathbf{s}_{ij}^\mathrm{T} \mathbf{a}_{ij}$, where $\mathbf{a}_{ij}$ is the N_z-dimensional vector representing the slice of A_{ijk} in the z direction, and $\mathbf{s}_{ij}$ is the corresponding sensing vector. Stated this way, for each i, j,the above relation yields a single linear equation with N_z unknowns, or, said differently, it is an N_z-undercomplete system. Typically, N_z is in the range of hundreds, making the reconstruction challenging for any practical compressed sensing scheme. However, extra equations can be obtained by projecting a set of time-multiplexed patterns, or by increasing the voxel size, or a combination of the latter two. Given that the total number of equations obtained is noted by n, we can write

$$\mathbf{e} \;\approx\; \mathbf{S}\mathbf{a}, \quad (6)$$

where $\mathbf{S} \in \mathbb{R}^{n \times N_z}$ is referred to as the sensing matrix. we henceforth denote the *compression ratio* of our reconstruction problem by $\kappa = \frac{n}{N_\mathrm{z}}$ referring to the latter quantity as the *compression ratio* of the measurements. The solution of such reconstruction problem falls within the domain of compressed sensing [8], constraining the sensing matrix $\mathbf{S}$ to satisfy the restricted isometry property (RIP [9]).

Input: Measurements **e**, sensing matrix $\mathbf{S}$, analysis dictionary $\mathbf{\Omega}$, weight λ, parameter $\tau > 0$
Output: Data **a**
Initialisation: $\boldsymbol{\mu}^0 = \mathbf{0}, \mathbf{z}^0 = \mathbf{0}$
for $k = 1, 2, ...$*until convergence* **do**

$$\mathbf{a}^{k+1} = \left(\mathbf{S}^\mathrm{T}\mathbf{S} + \tau \mathbf{\Omega}^\mathrm{T}\mathbf{\Omega}\right)^{-1} \left(\mathbf{S}^\mathrm{T}\mathbf{e} + \tau \mathbf{\Omega}^\mathrm{T}(\mathbf{z}^k - \boldsymbol{\mu}^k)\right)$$

$$\mathbf{z}^{k+1} = \sigma_{\frac{\lambda}{\tau}}\left(\mathbf{\Omega}\mathbf{a}^{k+1} + \boldsymbol{\mu}^k\right)$$

$$\boldsymbol{\mu}^{k+1} = \boldsymbol{\mu}^k + \mathbf{\Omega}\mathbf{a}^{k+1} - \mathbf{z}^{k+1}$$

end

Algorithm 1: Alternating direction method of multipliers (ADMM, [12]). Here, $\sigma_t(x) = \mathrm{sign}(x)\max\{|x| - t, 0\}$ denotes the element-wise soft thresholding.

3. SIGNAL RECONSTRUCTION

Since equation (6) represents an undercomplete system, prior information on $\mathbf{a}_k$ has to be added. One of a commonly used priors, usually referred to as *sparse analysis* is that the reconstructed signal a admits a sparse projection on the analysis dictionary $\mathbf{\Omega}$. Reconstruction of the scattering coefficient a from equation (6 can be cast as the optimization problem

$$\min_{\mathbf{a}} \|\mathbf{\Omega}\mathbf{a}\|_1 \quad \text{s.t.} \quad \|\mathbf{S}\mathbf{a} - \mathbf{e}\|_2 \le \epsilon, \quad (7)$$

where ϵ represents the amount of tolerable noise.

In this work, the analysis dictionary $\mathbf{\Omega}$ was selected as the concatenation of the identity matrix with the matrix of derivatives following [10]. This choice promotes the solution to comprise a sparse set of piece-wise constant regions.

3.1. Pursuit via ADMM

Problem (7) is non-smooth due to the ℓ_1 norm in the objective; however, it can be solved using the standard proximal projection methods through *Bregmann splitting* [11]. The splitting introduces an auxiliary variable $\mathbf{z} = \mathbf{\Omega}\mathbf{a}$; the resulting problem is typically rewritten as the unconstrained minimization problem of the augmented Lagrangian

$$\min_{\mathbf{a},\mathbf{z}} \frac{1}{2}\|\mathbf{S}\mathbf{a} - \mathbf{e}\|_2^2 + \frac{\tau}{2}\|\mathbf{\Omega}\mathbf{a} - \mathbf{z} + \frac{1}{\tau}\boldsymbol{\mu}\|_2^2 + \lambda\|\mathbf{z}\|_1 \quad (8)$$

Here, the parameter τ controls the penalty strength, and $\boldsymbol{\mu}$ is the vector of Lagrangian multiplies associated to the equality constraint $\mathbf{z} = \mathbf{\Omega}\mathbf{a}$. The combination of Bregmann splitting with augmented Lagrangian is known as the *alternating direction method of multipliers* (ADMM), summarized as Algorithm 1.

3.2. Fast approximation via ADMM networks

Recent work by [3] followed by [4, 5, 6, 7] advocate for an alternative approach of exploiting sparse priors. Coming from the classical machine learning background, [3] showed that the popular iterative shrinkage-thresholding algorithm (ISTA [13]) from the family of proximal projection methods can be truncated after a small number of iterations and unrolled into a feed-forward architecture resembling a neural network, with the parameters of the layers dictated by the ISTA iterations. In [4, 5, 6, 7], the authors proposed to apply a similar idea to the ADMM algorithm to approximate the solution of analysis models similar to problem (8). Following this methodology, Algorithm 1 is converted into a fixed-depth feed-forward network (henceforth, *ADMM network*). The network is then trained supervisedly to approximate the optimal solution.

Each iteration of Algorithm 1 can be viewed as a layer of the network, receiving $\mathbf{z}^k$ and $\boldsymbol{\mu}^k$ as the inputs, and producing $\mathbf{z}^{k+1}$ and $\boldsymbol{\mu}^{k+1}$ as the outputs, with k denoting the layer number. The only exception is the output layer $k = K$, producing the (approximate) solution $\mathbf{a}$.

The training of the network is performed on a collection of pairs of training vectors, comprising a ground truth signal $\mathbf{a}^*$ and its measurements $\mathbf{e} = \mathbf{S}\mathbf{a}^*$. At training a loss function of the form

$$\mathcal{L} = \hat{\mathrm{E}}\|\mathbf{a}^* - \mathbf{a}(\mathbf{e})\|_2^2 \tag{9}$$

is minimized using standard stochastic gradient [14], where $\hat{\mathrm{E}}$ denotes the empirical ensemble average. We followed the stochastic gradient training schemes detailed in [6].

3.3. Supervised learning of the sensing matrix

Although there exist previous studies aiming at finding optimal sensing matrices (see, e.g., [15] and [16]), here we focus on *binary* matrices. Algorithm 2 summarizes the proposed learning scheme. At each iteration, a random set of elements in $\mathbf{S}$ is flipped (i.e., 0 becomes 1, and 1 becomes 0). The new matrix is used to produce the measurements $\mathbf{e} = \mathbf{S}\mathbf{a}^*$ on a training set comprising a collection of ground truth data vectors $\mathbf{a}^*$. The reconstruction algorithm is applied, producing $\hat{\mathbf{a}} = \mathbf{a}(\mathbf{e})$. If the loss $\epsilon = \hat{\mathbb{E}}\|\mathbf{a}^* - \hat{\mathbf{a}}\|_2^2$ is reduced, the new matrix is retained. It is straightforward to show that the proposed algorithm produces a monotonically converging sequence of errors $\{\epsilon_k\}$.

Input: $\mathbf{a}, \mathbf{S}_0, \boldsymbol{\Omega}$
Output: $\mathbf{S}$
Initialisation:
$\mathbf{S} = \mathbf{S}_0$, $\hat{\mathbf{a}} = \arg\min_{\mathbf{z}} \frac{1}{2}\|\mathbf{S}\mathbf{a} - \mathbf{S}\mathbf{z}\|_2^2 + \lambda\|\boldsymbol{\Omega}\mathbf{z}\|_1$,
$\epsilon_0 = \sum \|\hat{\mathbf{a}} - \mathbf{a}\|_F^2$
for $k = 1, 2, ...$ *until convergence* **do**
  Select p at random from $\{1, \ldots, m \times n\}$
  Set $\mathbf{S}_c = \mathbf{S}$ and flip p random dimensions in $\mathbf{S}_c$
  $\hat{\mathbf{a}} = \arg\min_{\mathbf{z}} \frac{1}{2}\|\mathbf{S}_c\mathbf{a} - \mathbf{S}_c\mathbf{z}\|_2^2 + \lambda\|\boldsymbol{\Omega}\mathbf{z}\|_1$
  if $\epsilon_k > \sum \|\hat{\mathbf{a}} - \mathbf{a}\|_F^2$ **then**
    $\epsilon_k = \sum \|\hat{\mathbf{a}} - \mathbf{a}\|_F^2$;
    $\mathbf{S} = \mathbf{S}_c$
  else
    $\epsilon_k = \epsilon_{k-1}$;
end

Algorithm 2: Greedy learning procedure for a binary sensing matrix.

4. COMPRESSED REGISTRATION

Traditionally, compressed sensing approaches focus on the *reconstruction* of a latent signal from a set of compressed measurements. In some applications, however, a reasonable knowledge is available about the signal being reconstructed, while its position or deformation is unknown. We are therefore interested in the problem of *compressed registration*. We assume the volume $\mathbf{A}$ being imaged $\mathbf{A}$ to be created through an unknown transformation T from some parametric group of transformations $\mathcal{T}$ of a known volume $\mathbf{A}^*$. The goal of compressed registration is to estimate the transformation T given the measurements $\mathbf{E}$ of $\mathbf{A}$. Formally, the problem can be cast as the minimization

$$\min_{T \in \mathcal{T}} \|\mathbf{E} - \mathbf{S}_T(\mathbf{A})\|_{\mathrm{F}}^2. \tag{10}$$

The template volume $\mathbf{A}^*$ is assumed to be given in sampled form on a fixed grid $\mathcal{G}^*$, while the sensing matrix $\mathcal{G}$ operates on a possibly different grid $\mathcal{G}$.

Here, as a proof-of-concept, we restricted $\mathcal{T}$ to be the three-dimensional translation group.

5. RESULTS

We present volumetric image reconstruction results on a prototype acquisition setup depicted in Figure 2. We used a computer-controlled DLP Lightcommander projector to illuminate the scene, and a FireWire Point Grey Flea2 camera to capture the images. Scattering models were created by engraving small dots in a clear glass volume using a commercial laser engraving process. Acquisition was performed on a 70x60x50 mm volume by projecting between 4 to 16 128x128 images, captured by a 1280x960 pixels monochrome camera. Reconstruction was executed on an Intel 2600Mhz quad-core computer. This acquisition environment simulates to some extent the physics of fan beam x-ray scattering tomography, with the important difference that x-rays scattering is dominated by Compton scattering, while in our visible light system Mie (and, possibly, Raleigh) scattering is dominant [17].

In what follows, we report the reconstruction results from measurements with different compression ratios and different reconstruction methods. The networks were trained using data containing shapes geometrically similar to the model. Different number of layers (4, 8, 16, 25 and 32) was used

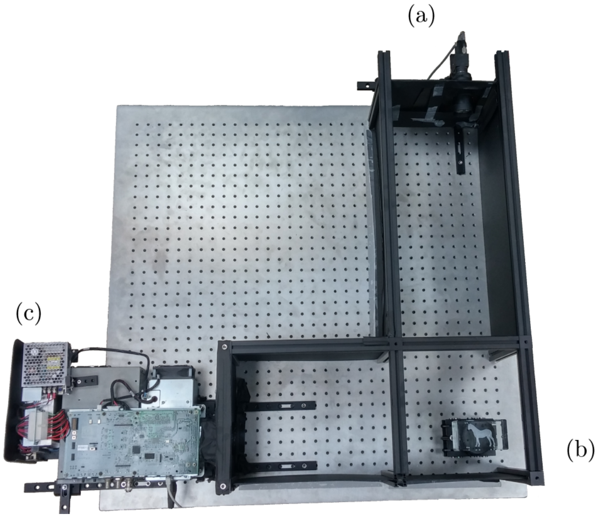

Fig. 2. Top view of the experimental acquisition setup showing the camera (a), model (b) and projector (c).

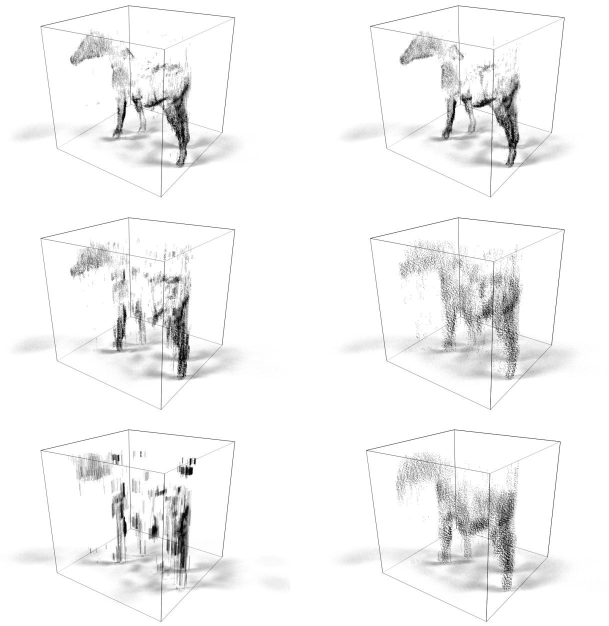

Fig. 3. Reconstruction results using ADMM on the acquired *Horse* models for random (left column) and optimal (right column) binary projections with compression ratios values of $1:8$ (top row), $1:16$ (middle row), and $1:32$ (bottom row).

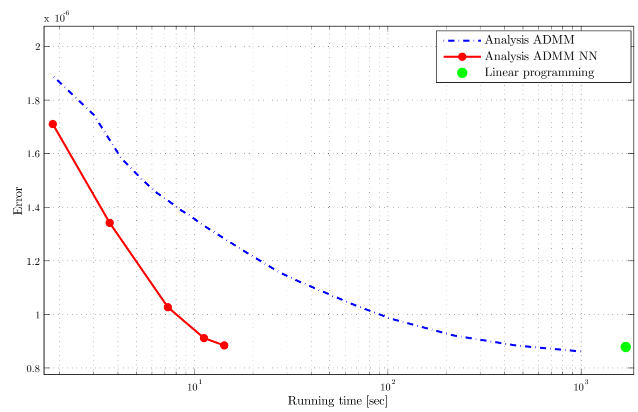

Fig. 4. Comparison of reconstruction performance of the exact iterative ADMM and its approximation by ADMM networks with different number of layers on the acquired *Horse* model with $\kappa = 1:8$.

as described in Section 3.2. Figure 3 depicts the reconstruction results with a random sensing matrix and the optimal matrix trained using Algorithm 2. Figure 4 compares the reconstruction accuracy as function of execution time for the exact ADMM algorithm and its approximation via the ADMM network. As a reference, we show the performance of the standard basis pursuit algorithm implemented as linear programming [10].

Finally, we observed that the non-reconstructive compressed registration task succeeds with a single projection; the quality of *reconstruction* from such highly undetermined data is clearly unacceptable.

6. CONCLUSIONS

In this research, we studied the problem of reconstructing volumetric images by means of scattering tomography. The proposed techniques were evaluated on real data obtained from a custom-built visible light acquisition system. Our results demonstrated accurate reconstruction from measurements with ten-fold compression ratios; these results improve significantly when optimal illumination patterns are used in lieu of random ones. Furthermore, we showed that a significantly better compression can be obtained in non-reconstructive tasks such as compressed registration, often achieving faithful alignment from a single projection only. We also evaluated the running times of the proposed iterative reconstruction algorithm and its fast approximation. Although our Matlab implementation is far from being optimal, this evaluation suggests at least an order of magnitude improvement in runtime at comparable reconstruction accuracy. Though the presented models and experiments assumed visible light imaging, we believe that the presented methodology is suitable for x-ray scattering tomography. The relatively small amount of measurements and the fast reconstruction numerics promise potential applications in real-time medical imaging, including in particular 3D vascular imaging. We intend to explore these applications in future research.

7. REFERENCES

[1] W. Cong and G. Wang, "X-ray scattering tomography for biological applications," *Journal of X-Ray Science and Technology*, vol. 19, no. 2, pp. 219–227, 2011.

[2] S. J. Norton, "Compton scattering tomography," *Journal of applied physics*, vol. 76, no. 4, pp. 2007–2015, 1994.

[3] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. ICML*, 2010, pp. 399–406.

[4] P. Sprechmann, A. M. Bronstein, and G. Sapiro, "Learning robust low-rank representations," *arXiv preprint arXiv:1209.6393*, 2012.

[5] P. Sprechmann, A. M. Bronstein, and G. Sapiro, "Real-time online singing voice separation from monaural recordings using robust low-rank modeling.," in *ISMIR*, 2012, pp. 67–72.

[6] P. Sprechmann, R. Litman, T. Ben Yakar, A. Bronstein, and G. Sapiro, "Efficient supervised sparse analysis and synthesis operators," in *Proc. NIPS*, 2013.

[7] T. Ben Yakar, R. Litman, P. Sprechmann, A. Bronstein, and G. Sapiro, "Bilevel sparse models for polyphonic music transcription," in *Proc. ISMIR.*, 2013.

[8] D. L. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[9] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[10] J. Gu, S. K. Nayar, E. Grinspun, P. N. Belhumeur, and R. Ramamoorthi, "Compressive Structured Light for Recovering Inhomogeneous Participating Media," in *European Conference on Computer Vision (ECCV)*, Oct 2008.

[11] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[12] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[13] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on pure and applied mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[14] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010*, pp. 177–186. Springer, 2010.

[15] J. Mairal, F. Bach, and J. Ponce, "Task-driven dictionary learning," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 4, pp. 791–804, 2012.

[16] R. Rubinstein, T. Peleg, and M. Elad, "Analysis k-svd: A dictionary-learning algorithm for the analysis sparse model," *Signal Processing, IEEE Transactions*, vol. 61, pp. 661 – 677, 2000.

[17] C. F. Bohren and D. R. Huffman, *Absorption and scattering of light by small particles*, Wiley. com, 2008.